DOCUMENT RESUME

ED 117 944 FL 006 311

AUTHOR Greenberg, Joseph H.
TITLE Numeral Classifiers and Substantival Number: Problems in the Genesis of a Linguistic Type. Working Papers on Language Universals, No. 9.
INSTITUTION Stanford Univ., Calif. Committee on Linguistics.
PUB DATE 72
NOTE 39p.; Revised version of a paper presented at the International Congress of Linguistics (11th, Bologna, Italy, 1972)

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage
DESCRIPTORS Classical Languages; Comparative Analysis; *Contrastive Linguistics; Diachronic Linguistics; Grammar; Language Patterns; Language Research; Language Typology; *Language Universals; Linguistic Theory; Nominals; Number Concepts; *Numbers; *Number Systems; Synchronic Linguistics; *Syntax
IDENTIFIERS Count Nouns; Mass Nouns

ABSTRACT

This study is based on a sample of about 100 languages with numeral classifiers. An attempt is made at reconstructing the dynamics of the process by which such systems arise, develop, and decay. Among the hypotheses advanced are the following: (1) numeral classifiers involve the overt expression of one kind of quantification, namely, counting by units; (2) the numeral classifier construction is modelled after the measure construction with mass nouns and hence arises in languages with previous mass-count noun distinction; and (3) the classifier in a numeral classifier language has the same function as a singulative does in a language with a collective-singulative distinction. (Author)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED117944

Working Papers on Language Universals
No. 9, November 1972
pp. 1-3

RESEARCH PAPERS

NUMERAL CLASSIFIERS AND SUBSTANTIVAL NUMBER: PROBLEMS IN THE GENESIS OF A LINGUISTIC TYPE*

Joseph H. Greenberg
Stanford University

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
JOSEPH H. GREENBERG
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

FL006311

ABSTRACT

The present study is based on a sample of about 100 languages with numeral classifiers. An attempt is made at reconstructing the dynamics of the process by which such systems arise, develop, and decay. Among the hypotheses advanced are the following: 1. Numeral classifiers involve the overt expression of one kind of quantification, namely, counting by units. 2. The numeral classifier construction is modelled after the measure construction with mass nouns and hence arises in languages with previous mass-count noun distinction. 3. The classifier in a numeral classifier language has the same function as a singulative does in a language with a collective-singulative distinction.

---

* This is a slightly revised version of the paper presented to the Eleventh International Congress of Linguists in Bologna (1972) at the plenary session on Universals. An expanded and more thoroughly revised version is being prepared for a collection of papers from the Bologna congress to be edited by A. Makkai and to be called Linguistics at the Crossroads.

The data and hypotheses presented here are part of a broader investigation concerned with numeral classifier systems considered both as representatives of a type of nominal classification and in relation to the problems of quantification in language.[1] What is meant by quantification in this connection is the manner in which languages express the fact that reference is being made to a quantitatively delimited amount of the thing mentioned.[2]

Such a typological approach involves both synchronic and diachronic considerations. Initially we take into an account an extensive, ideally, an exhaustive sample of languages which is based on preliminary notions regarding the definitional characteristics of the type. A comparison of such languages leads to a number of synchronic generalizations, usually implicational in form. The second major aim is to uncover the dynamic principles, that is the recurrent types of change in historically independent instances involved in the rise, subsequent expansion and ultimate dissolution of the type.[3] In carrying out this part of the investigation our methods include deductions based on internal reconstruction within individual languages, the comparative method within linguistic stocks and direct historical documentation where this is available.

As noted initially, the tentative conclusions presented here are but a portion of a broader study which is in progress. In the present study the emphasis will be on questions relating to the initial conditions under which numeral classifier systems may be conjectured to arise. In the final section, in order to place the present study within the more general perspective of the study as a whole, a series of other problems and in some cases hypotheses regarding them will be outlined without pursuing them in detail.

---

[1] The present research was supported by the National Science Foundation as part of the Language Universals Project at Stanford University.

[2] The somewhat vague term 'thing mentioned' is used here because although the present study is basically confined to nominal phrases, verbal action can also be quantified. This is briefly discussed in the final section of the paper.

[3] For discussion and exemplification of these methods, see Greenberg 1966, 1969, 1970(a), 1970(b).

One limitation should be mentioned at the outset. Systems of the type with which we are concerned here have undoubtedly, in some instances, arisen under conditions of language contact. For example, in those DRAVIDIAN languages which have such systems it seems clear that they have developed in general as a result of contact with INDO-ARYAN languages.[4] Any theory of origin will ultimately have to take into account both the conditions under which pristine systems arise and those in which contact is a major factor. In the study in the present form, the predominant emphasis is on the former.

As mentioned earlier the sample is not exhaustive and this, of course, adds still further to the tentative nature of the results. Nevertheless, the hypotheses presented here are based on quite extensive data.[5] They are presented here, in the hope that they may provide at least a basis for conclusions that can be tested and modified in the light of both raw data and more penetrating theoretical analysis.

We begin with an attempt at a preliminary definition of what constitutes a numeral classifier language in terms of the existence of a particular

---

[4] On this topic see especially Emeneau (1956).

[5] A list of languages in my samples follows. In a few instances the numeral classifier system is very marginal, e.g. BULGARIAN, because of the use of duši 'soul' used in enumerating persons, and HUNGARIAN, because it has a numeral series used only with persons: AHOM, AINU, ASSAMESE, BANGGAIS, BENGALI, BLACK THAI, BODO, BRETON (MEDIEVAL), BRIBRI, BROU, BULGARIAN, BURMESE, CEBUANO, CHINESE (ARCHAIC, MANDARIN, HAKKA, CANTONESE), CHIRIPO, CHOLON, CHONTAL (MAYAN), CUNA, DAY, DIOI, ENGENNI, EGYPTIAN ARABIC, EMPEO, FIJIAN, GARO, GILBERTESE, GILYAK, GUAYMI, HAIDA, HAUSA, HUNGARIAN, HUPA, IBAN, IBIBIO, IRISH, ISHKASHIM, JACALTEC, JAPANESE, KACHIN, KAREN, KARO-BATAK, KATU, KEI, KHAMTI, KHARIYA, KHASI, KHMER, KHMU, KIRIWINA, KOLAMI, KOREAN, KURUKH, LAOTIAN, LISU, MALAY, MAN, MARU, MERIR, MIKIR, MIRI, MON, MOTA, MUCHIK, NAHUATL (CLASSICAL and TETELCINGO), NAURU, OJIBWA, OMANI ARABIC, OSSETE, PALAU, PALAUNG, PARJI, PASHTO, PERSIAN, POCOMCHI, PONAPE, PUR, RAWANG, SAMOAN, SHAN, SONSOROL, TAJIK, TARAON, TARASCAN, TAT, THAI, THO, TLINGIT, TOBA-BATAK, TOTONAC, TRUKESE, TSIMSHIAN, TURKISH, TZELTAL, TZOTZIL, UVEA, UZBEK, VIETNAMESE, WHITE THAI, WOLIO, YUROK.

syntactic construction. A considerable number of the world's languages including almost all of these in Southeast Asia exhibit the following characteristic. An ENGLISH phrase such as 'five books' is rendered in translation by a phrase containing, outside of possible grammatical markers, not two but three elements. The kind of literal translation often supplied in grammars of such languages might be something like 'five flat-object book'.[6] The second item in such a phrase is often called a numeral classifier in allusion both to its occurrence in a numeral phrase and to its providing a semantic classification of the head noun.

Implicit in the terminology 'numeral classifier', there is, then, a quite straightforward definition of the syntactic construction in which the classifier appears, the occurrence of which could be criterial for a language to be considered a numeral classifier language. For example, the following statement by Burling (1965, 244) might provide the basis for a definition along these lines; "In many languages of Southeast Asia, a number is never used without being accompanied by one of the special morphemes known as classifiers."

However while a useful starting point for discussion, it is clear that simply rephrasing this statement as a definition would leave unsolved a number of questions, questions which require settlement before a definition can be applied. This becomes particularly obvious when confronted with the variabilities and complexities of languages usually assigned to the type. For example, such a definition might be interpreted to require that every noun which in a language like ENGLISH may be preceded by a number, should, in a classifier language, have a classifier. On such a view it is not excessive to state that there are no numeral classifier languages. There are, in fact, particular classes of nouns, e.g. measures, units of time, and the word 'time' in such phrases as 'three times' which hardly ever occur with classifiers. In some languages, always considered to be numeral classifier languages, the group of nouns which do not take classifiers is still more extensive (e.g. VIETNAMESE).

---

[6] Of course, other word orders are possible.

In many languages the classifiers are not compulsory even for the restricted set of nouns that have them. This holds for example in KHMER in which, we are told, the expression without the classifier is stylistically less formal.[7]

Sometimes the restrictions on the classifier construction pertain to the numbers with which the classifiers may cooccur. For example, in KHASI and TAT they do not appear within the number 'one', while in MALTO they only occur with numbers larger than two, in this case with the numbers borrowed along with the classifier system from INDO-ARYAN. It is particularly common for classifiers not to occur with higher units of the numerical system and their multiples e.g. 10, 20, 60, 100, 300.

Syntactically, also, there is variability in that the classifiers need not be confined to numerical constructions. In MANDARIN and other languages the classifier is required with demonstrative even in non-numeral phrases. In other languages it may occur in such phrases usually with some difference in meaning between instances in which the classifier does or does not appear (e.g. THAI). In THAI, it may also occur with qualifying adjectives under the same general circumstances as with demonstratives. In KIRIWINA it is required with demonstratives and certain adjectives while it may not occur with certain other adjectives. In some languages it may occur with the noun in the absence of any modifiers, numeral or otherwise, and is thus a kind of article (e.g. in DIOI, a THAI language). In one MAYAN language, at least, JACALTEC, the classifier can occur as the sole constituent of a substantive phrase in its function as an anaphoric substitute i.e. as a pronoun. It is indeed universal in languages with numeral classifier constructions that the head noun may be deleted either when it has been either previously mentioned or can be supplied from the non-linguistic context. In instances like DIOI and JACALTEC we may legitimately ask whether, synchronically considered, these systems should be considered numeral classifier as distinct from some other type of nominal classification.[8]

---

[7] According to Jacob (1965:148).

[8] Diachronically, from the evidence of related languages, they have in all probability arisen from systems in which the classifiers were confined to numeral constructions.

If all this creates difficulties in establishing precise criteria based on syntactic function in numerical phrases, even graver problems arise in regard to classification as a definitional criterion. It is clear that many of the items that are listed as classifiers in grammars of numeral classifier languages cannot, on any reasonable view, be said to classify. In our initial example, we employed a gloss of the type frequently found in grammars of such languages, namely, 'five flat-object book'. Taking such a translation at its face value, we can justifiably state that we have classification in the semantic sense because, indeed, a book is a kind of flat object. The word for 'tail' is sometimes used as a classifier for animals (e.g. ekor in MALAY) but we cannot consider a dog a kind of tail though of course we can devise a property 'having a tail'. On the other hand we could define the class meaning of ekor in MALAY as that which is common to all nouns which take ekor as a classifier. This is, of course, what is usually done in describing class meanings in noun-class languages. These two alternatives bring out an interesting difference between numeral classifier and the noun-class language with which they have sometimes been compared. In the former, in the majority of instances, the classifier is itself a noun with its own lexical meaning and may, in fact, have its own classifier when it functions as the head of a noun phrase.

But even the approach based on the meaning of the items with which a classifier cooccurs and which disregards the lexical meaning of the classifier itself, runs into difficulties which are similar to those incidental to the establishment of class meanings in noun-class languages. For example in THAI tua is used with animals in general but it also occurs with other nouns e.g. sʊ̂a 'a coat'.[9]

Furthermore, in some languages such as BURMESE and THAI, there are a fair number of words which are, as it were, their own classifiers.

---

[9] Noss (1964:106) seeks to define tua as occurring with non-humans with anthropomorphic characteristics, e.g. animals, coats, trousers, and tables, the last three because they have arms or legs.

An example is in BURMESE ʔein ta-ʔein "house one-house" in which ʔein in its first occurrence is a head noun and in its second occurrence a "classifier". In this limiting case the approach through the lexical meaning of the classifier and the semantic properties of the cooccurring noun fall together but give the somewhat fatuous result that ʔein as a classifier means 'the property of being a house'.

At the other end of the logical spectrum there are examples like buʔuk 'piece' in CEBUANO, a language of the Philippines, which is used with any classifiable noun so that we would have to assign it the meaning 'having the property of being a classified noun'.

All this does not, of course, destroy the notion that, in a purely formal sense, common cooccurrence with the same classifier determines a classification of those head nouns which occur with classifiers even though such a classification is often formal rather than semantic, is non-exhaustive in relation to the nouns of the language, is frequently overlapping in that the same noun occurs with more than one classifier and that classes with one member (BURMESE) or overall classifications with only one class (CEBUANO) may be found.

The considerations just cited in regard to classification have, in fact, been widely appreciated so that many who have been engaged in the description of these languages have consciously eschewed the term "classifier" in favor of some semantically more neutral term. Such terms are legion e.g. numerical coefficient (Anceaux, WOLIO); numeral adjunct (Fraser, LISU); numerical determinative (Milner, PALAUNG), while in the RUSSIAN literature numerativ has been widely adopted for this purpose.

The foregoing considerations might be held to destroy the very notion of languages with numerical classifiers as a valid linguistic type. Nevertheless, there is still an important difference between languages which are generally held to belong to this type and those which are not, although our initial discussion has failed to capture it.

In order to isolate what is distinctive about these languages, we may first consider a range of facts which have not yet engaged our attention. In general, grammars of such languages as BURMESE, THAI and MANDARIN subsume under the same basic construction of numeral + classifier + noun not only such examples as five + flat-object + book, but also many others in which, in contradistinction from this one, the item corresponding to the classifier requires translation into languages like ENGLISH. Moreover, in languages like KHMER for which it is stated as the general rule that classifiers are optional in these other instances classifiers are not optional.[10]

The most important class of such instances is probably the measure construction which occurs most characteristically with mass nouns.[11] In non-classifier languages in which a grammatical mass/count distinction exists, a central characteristic of mass nouns is that they normally do not enter into a direct construction with a numeral but require an intervening measure e.g. 'one cup of water', 'two gallons of water'.

What has impressed students of languages such as THAI is the evident parallelism between such expressions as:

1. ka'fɛ· sɔ̌ŋ thûaj "coffee two cup"
2. bùrì sɔ̌ŋ muan "cigarette two long-object"

Most linguists who have described these languages have felt that these are at best subtypes of the same fundamental construction. Most commonly they have used some common expression for all constructions of this same general type and then distinguished a series of subtypes, one of which is exemplified by the second of the above phrases.

In the foregoing examples the contrast was between a measure and a count construction. The following set of contrasts, once more from THAI, will show that, as exemplified by the first two constructions, certain count constructions display the same property as measure constructions in that

---

[10] Jacob (1965:145).

[11] Of course, measures of weight can occur with all kinds of physical objects including countables.

the elements in the classifier position require translation into ENGLISH and cannot be dispensed with in languages like KHMER.

1. bùrî sɔ̌ŋ sɤŋ "cigarette two pack"
2. bùrî sɔ̌ŋ lɔ̌ "cigarette two dozen"
3. bùrî sɔ̌ŋ muan "cigarette two long-object"

Evidently what is peculiar to languages like THAI is the overt expression of one particular mode of quantification, namely counting by units. This manner of quantifying is evidently the "unmarked" method in that, in the absence of an overt indication, unit counting is assumed.[12]

The point noted here is not novel. Some analysts have understood its special status and have employed such expressions as "unit counters" or "individual classifiers" to mark out this particular class of expressions.[13] If we reserve the term "classifier" for such unit counters we may now, in a closer approximation (but nevertheless as we shall see later only an approximation) to the definition of the characteristic numeral classifier construction, delimit it in terms of the overt expression of unit counting.

We may say then, that in even the most elaborate system, all the classifiers are from the referential point of view merely so many ways of saying 'one' or, more accurately 'times one'. The latter expression is to be preferred because, taken pragmatically, there is a difference between numerals proper and modes of quantifying even when the latter involve a number. 'Two dozen' and 'twelve pairs' represent different kinds of quantification acts even though the identity of the final numerical result is guaranteed by the commutative law of multiplications. Hence unit counting is to be distinguished from 'one' as a numeral although the connection between the two is a close one.

---

[12] I have encountered just one instance of non-unit counting as the preferred form. According to Bataillon (1932:10), in UVEA counting is most commonly by twos. There are unit classifiers when counting by units is intended. Even here, however, counting by twos also requires an overt indicator, e.g. ufi lua gafua "yam two classifier" 'two yams'; ufi lua gahoa "yam two pair" 'four yams'.

[13] The earliest statement along these lines that I have encountered is that of Emeneau (1951: 53) who gives as the class meaning of classifiers as "one unit quantity or number of that denoted by the noun it precedes". Note, however, that measures are included here (i.e. "quantity").

It was noted earlier that analysts of numeral classifier languages have often felt that unit counters in these languages are essentially no different from other quantifiers in these languages, that the difference is not intrinsic to the language being analyzed but is imported from considerations of differences of translation into languages not belonging to this type.

However, there is some evidence that both formally and psychologically the unit-counter is a unique type in these languages even though it is, in general, affiliated with the more inclusive type. For example, Chao, at once an eminent linguist and a native MANDARIN speaker, classifies unit counters as a special type of measure under the term "individual classifier" and notes that they do not cooccur with mass nouns (Chao, 1968:503). Burling reports in regard to the most common unit counter of BURMESE (1965:262), the so-called general classifier -khù, that it is included by some BURMESE speakers "in the same series as the classifiers for the powers of ten... -khù indicating only one individual object". It was noted earlier that multiples of higher numerical units often do not take classifiers. This also occurs in BURMESE and shows clearly the function of the unit classifier as meaning 'times one'. Thus in BURMESE "two-ten book" = 20 books, i.e. 2 x 10 books while, following the interpretation by native speakers just cited "two-khù (classifier) book" '2 books, i.e. 2 x 1 'books'. Many analysts consider words for 'ten', 'hundred' etc. in these languages as a subtype of classifiers.

In a few languages there are grammatical peculiarities which distinguish count from measure constructions without there usually being sufficient information to decide whether this separates unit counting from all other types of quantification or simply counting as against measuring. For example measures take a different linking particle in CEBUANO.

In spite of these few instances, the overwhelming impression is that of at least surface conformity of all quantifying constructions in these languages in such matters as word order and syntactic markers. This is so much the case that the first and simplest general diachronic hypothesis

would be that they have modelled the unit counting construction after preexistent measure and non-unit count constructions.[14]

Nevertheless, the definition of a numeral classifier language as one which contains a construction in which counting by units receives overt expression raises some further problems. Such a formulation, since it primarily defines a construction and only indirectly a language type taken by itself leaves unresolved such questions as whether a language in which the classifier is always optional or in which there are only isolated instances of it, e.g. ENGLISH, because of expressions like 'two head of cattle', is to be considered a numeral classifier language. From the dynamic point of view, however, this is not our major concern which is rather the genesis, spread and loss of such constructions within languages.

There remains still another problem and this pertains to the adequacy of the definition of the construction itself as one in which unit counting receives overt expression. This stands in need of further elucidation regarding what is to be considered a unit counter. When writers of grammars in ENGLISH seek to explain the notion of numeral classifier their stock example is 'head of cattle'. Yet other types of expressions occur in ENGLISH which, it might be argued involve unit counters, for example, 'sheet of paper'. The existence of a contrast between 'sheet of paper' and ream of paper', the latter being defined as equivalent to 480 sheets, seems to suggest the status of 'sheet' as a unit counter. Similar considerations hold regarding expressions such as 'slice of bread', 'piece of meat' and many other. They are countable and contrast as units to such non-unit counters as 'bunch' in 'bunch of carrots'. Yet their presence in ENGLISH and in many other languages is not, in itself, generally considered a basis for considering the language a numeral classifier language. On the other hand, the nouns themselves in ENGLISH are grammatically mass nouns in these constructions, but so is 'cattle'.

---

[14] The same basic hypothesis seems to be stated in Sen-Gupta (1970, especially 677-8) as indicated in the following remark, "We consider MW [i.e. measure word] as the basis of NuCl [i.e. numeral classifier]".

Is there any basis for this intuitive feeling? It may be proposed that it is a certain arbitrariness as to what constitutes an individual in such instances which underlies this reaction, an arbitrariness not present in such instances as 'one dog' or 'one automobile'. A homely conceptual experiment may serve to pinpoint this difference. If I cut a piece of meat in two, I have two pieces of meat, but if I cut a dog in two, I still have only one dog, a dead one. The property that distinguishes dogs and automobiles in these cases is evidently internal organization into an integrated and organic whole, whether natural in the case of the dog or artificial in the case of the automobile. We might call this feature ± structured.

There is still another kind of borderline case which can be illustrated by ENGLISH phrases such as 'grain of sand', 'blade of grass', and 'strand of hair'. Once more we see types of phrases which are widespread in non-classifier languages and which could not in themselves lead one to classify them as numeral classifier languages. They would also not be employed as pedagogical examples in languages like ENGLISH in order to exemplify the notion of numeral classifier. Yet, as with the instances in the preceding paragraph they are unit items and countable. Moreover, they are, as it were, given in nature and do not have the same arbitrariness as 'piece' in 'piece of meat'. For this class of counters what makes them untypical is, it may be conjectured, their smallness and lack of individuality so that they are almost never used in actual counting. In this respect, the superordinates e.g. rice, grass, approach the status of liquids and other items which form the basis for the grammatical category of mass nouns. These "particulates" as we might call them are almost exclusively used with 'one' or the indefinite article and particularly frequent in negative statements e.g. 'In many stretches of the Sahara you will not find even one blade of grass.'. In this respect their quantification is frequently like that of mass items such as 'water' in that in non-measured constructions the universe of numeration is confined to an opposition between 'one' or 'a' and 'none', in construction with items

such as 'bit' or such other indefinites as 'some'.[15]

The two types of counters just considered have an equivocal status in that even analysts who have come to apprehend the difference among unit counters, non-unit counters and measures and the special status of the first of these in relation to the concept of numeral classifier, find difficulty in deciding in specific instances which of these morphemes are to be considered unit classifiers.[16] The reason for this uncertainty in practice is not merely that the analysis of the preceding section has not been carried out but because in many instances the same classifier has both "true" unit uses as well as the marginal meanings which would be excluded from consideration as numeral classifier constructions in a non-classifier language. For example, MANDARIN chang stated generally to be classifier for fairly extensive flat objects is used with 'paper' where it is to be translated as 'sheet' and with 'table' where there is no ENGLISH translation equivalent involved. Words like 'grain' are widely used both in the meaning of small particle and as a true classifier for small round or even large round objects. Indeed, these two classes of borderline items play a prominent role as sources for true classifiers in the course of the dynamic development of such systems. Where necessary, the two foregoing types will be referred to as "quasi-unit counters".

For purposes of the present study the terminology "unit counter" will continue to be employed with the understanding that quasi-unit counters are not included.

The discussion up to this point indicates that the languages commonly called numeral classifier languages can be considered from two points of

---

[15] In some descriptions such words are considered classifiers of mass nouns. But then they have the peculiarity that they can only occur with one. An example is BENGALI ṭuku as in ek-ṭuku jal 'a bit of water' as analyzed in Ferguson (1964).

[16] An instance in point is Winstedt (1945, 1957) who lists certain nouns as classifiers in his grammar of MALAY but not in his dictionary and vice versa. These seem generally to be quasi-unit classifiers e.g. potong 'slice'.

view, either as involving the overt expression of a particular mode of quantification, or as imposing a classification on the head nouns in numerical constructions. In effect, such languages may be said to belong to two typological classes simultaneously and, in this, there is no logical contradiction since anything can belong to more than one class simultaneously. Both of these approaches are legitimate and both are utilized in the broader study of which this is a part. However, the former of these will turn out to be the more relevant for questions of type genesis while the latter becomes valuable in considering further stages of dynamic development. In the light of our preliminary definition in terms of quantification we now turn to a more detailed consideration of the problem of type origin.

We might state our aim in terms of possible answers to the following question. What are the initial conditions in the form of other structural characteristics characteristic of languages which develop numeral classifiers? In putting our question in this form we are, at the most, asking for necessary, not sufficient conditions, that is, we are not in the position to assert that, given certain linguistic properties, a language will inevitably develop classifiers. This is surely beyond our powers. Even the more modest goal of necessary conditions, that is conditions which must be present for classifiers to arise might be too ambitious if by this we mean a single set of conditions. We may indeed have to deal with more than one type of origin. But this in turn should help understanding of these systems, since different origins usually imply different subtypological characteristics.

Our preliminary definition in terms of overt expression of unit counting and our observation regarding the virtually complete syntactic identity of all counting and measuring construction in these languages leads quite directly to a hypothesis which is negative in form. In spite of this it represents a kind of progress because it narrows the class of languages which have properties relevant to the rise of numeral classifier systems.

It is our working hypothesis that unit counters are modelled after the construction of mass nouns which cannot stand directly with numerals but require a measure or quasi-unit counter as an intermediary. Now such non-unit counters are found in virtually all the languages in the areas and linguistic stocks with which most linguists are concerned and it might be thought that they are universal. In fact, this seems not to be so. In particular there are a considerable number of AMERIND languages as well as some elsewhere, for example, in New Guinea which do not have measure constructions. Numerals occur directly both with nouns designating mass as well as countable objects. Hence no model exists in these languages for the development we postulate. Whorf (1941: 80) describes such a situation for HOPI. Unlike ENGLISH with its grammatical distinction of mass and count nouns, HOPI "has a formally distinguished class of nouns. But this class has no formally distinguished class of mass nouns.... One says not 'a glass of water' but kə·yí 'a water' ... not 'a piece of meat' but sik^{w}í 'a meat'".

There is evidently here a correlation between language and culture but not, I submit, in Whorf's terms of the philosophic non-existence of a Western world-view based on the Aristotelian dichotomy of form and matter. It is rather the absence of precise measures and their relatively infrequent employment which allows them to remain unexpressed since they can be deduced from context. This happens in special instances in languages like ENGLISH in restricted situations. Thus, in a restaurant, one can say "We want three coffees and one tea." and it will be understood that the unexpressed measures are 'cups'.

Although an evolutionary factor is involved here, it would be well to note in passing that the absence of a sufficient body of contrastive measures to require an explicit terminology is not correlated in any simple way with economic stages. While it is striking that MAYAN and NAHUATL figure among the AMERIND numeral-classifying languages we also find them in such languages as that of the YUROK of California who have no

domesticated plants or animals. The absence of agriculture does not preclude under certain circumstances a considerable accumulation of material goods and as in the case of the YUROK, the existance of a standard measure of value (dentalium shells).[17]

The foregoing hypothesis can also be stated as a synchronic implicational universal. The presence of unit counters implies the presence of measure and other non-unit counter type constructions.[18] To suppose then, diachronically, that there might be a development of unit counters in a language without overtly expressed non-unit counter types would be to hypothesize the genesis of what would be, if this is valid, a non-existent type.

We turn now to that vast majority of the world's languages which have measure constructions as well as various non-unit and quasi-unit count constructions. They therefore, by hypothesis, possess a model in accordance with which unit classifiers might come into existence. The question is whether there are any properties in addition to these which are relevant to such a development.

A number of synchronic generalizations can be made about numeral classifier languages some of which will be treated in due course in this paper. There is one, however, which is merely statistical, that is, has exceptions although it holds very widely. It will, I believe, shed light on the problem under discussion. Indeed it is precisely the study of the exceptions in this case which proves most useful.

Numeral classifier languages generally do not have compulsory expression of nominal plurality, but at most facultative expression. This has already been observed by Mary Sanches (1971) in an unpublished paper. She states her hypothesis as follows: "If a language includes in its basic mode of forming quantitative expressions numeral classifiers, then it will also have facultative expression of the plural. In other words it will not have obligatory marking of the plural on nouns." Sanches makes an additional valuable observation, namely, that the classified noun itself is normally singular. She includes in this such instances as ENGLISH

---

[17] The YUROK have been described as "primitive capitalists", cf. Goldschmidt (1951).

[18] By measure construction is meant here those of the type quantity + measure + noun. AMERIND languages which lack this construction still express measure by using a verb 'to measure' with a numeral, the precise kind of measure being deducible from context.

'cattle' in 'head of cattle'. As will appear in the discussion, it is advantageous to reinterpret this observation in the sense not of the singular but in that of lack of marking for number.

In addition to the handful of possible exceptions, some of them valid, others marginal, noted in her paper there are a few others from languages which happen not to figure in her sample e.g. OSSETIC and certain modern ARABIC dialects. Those latter are of particular interest since numerical classifiers do not occur in CLASSICAL ARABIC. Hence by studying these examples we can perhaps develop some insight into the conditions of their appearance. I have noted two fairly extensive systems in modern ARABIC dialects, in OMAN-ZANZIBAR and in EGYPTIAN ARABIC.[19] The present discussion is confined essentially to OMAN-ZANZIBAR which will hereafter be called OMANI.

In this form of ARABIC as described by Reichardt, a number of animals (described as schlachtbar), root crops, and the word for 'slave' are classified by ra:s 'head', a number of "horn-shaped" edibles by qarn 'horn', fruits by šo:b and flowers' by ʕo:d 'branch'. The system is therefore fairly extensive.

In OMANI, which has no case system, corresponding to the classical construction of 3-10 with the genitive plural, the plural is used with these numbers. With the other numbers except 'two' which employs either the dual or the number 'two' with the plural the singular is used, reflecting two classical constructions, one involving the accusative singular and the other the genitive singular.

In the OMANI classifying construction the numeral precedes the classifier, agrees with it in gender and governs it for number in accordance with the above rules. In this, the construction is entirely like that of a numeral with an unclassified noun. In the classifier construction the classified noun follows the classifier and is not affected by the syntax of the preceding construction. Examples with finda:l 'potato(es)', classified by ra:s 'head' will illustrate these rules. It should be noted that with

[19] For OMAN-ZANZIBAR ARABIC see Reichardt (1894) especially p. 85 and for EGYPTIAN ARABIC, Mitchell (1956: 94). Brockelmann (1908-13, II: 280), in addition to OMANI, gives an example from MALTESE in which ruḥ 'soul' is used in counting persons and 'head' for oxen and sheep in MODERN SYRIAC.

the singular and dual the numerals 'one' and 'two' are not commonly employed. Thus we have ra:s finda:l 'one potato, a potato', ra:se:n (dual of ra:s) finda:l 'two potatoes', thala:thit rwa:s (plural of ra:s) finda:l 'three potatoes', 9ashri:n ra:s finda:l 'twenty potatoes'.

What of finda:l which remains unchanged throughout? It is singular in form but is a collective and would in other contexts be translated 'potatoes'. In fact practically all of the words listed by Reichardt as taking this construction are collectives.[20] On the other hand not all collectives take this construction. The alternative corresponds to the use of the so-called 'noun of unity' (ʔismu ʾlwaḥdati) of CLASSICAL ARABIC. This is in CLASSICAL ARABIC a formation from the collective by the suffixation of the feminine suffix -at(un) (OMANI -a, -e) with a regular sound plural in a:t(un), OMANI a:t.[21] Where the collective has a noun of unity it is required both in CLASSICAL and OMANI ARABIC that it be used with numbers, the plural with 3-10 and the singular and the dual, usually without a numeral, for one and two respectively. An example of this alternative construction is OMANI be9u:d 'gnats' (coll.) which may not occur with a numeral. Based on the noun of unity be9u:da we have be9u:da 'one gnat'; thala:th be9u:da:t 'three gnats', etc. In at least one instance there is a choice of the two constructions. From baqar 'cattle' there is a noun of unity baqra 'a cow' but it can also be classified by ra:s in which case the collective, of course, is used. Thus 'one cow' is either

---

[20] Except 9abi:d 'slaves' which is an ordinary broken plural. There are several "psychological" parallels to this OMANI treatment of 'slave' as the only personal term with a classifier and in fact the same classifier as that used with animals. Vinogradov (1934: 94) states that in upper class RUSSIAN speech of the 18th century the collective numerals were used with words designating humans only when members of the lower social classes were involved. One would say dva arxiereja 'two archbishops' and not e.g. dvoje arxierejov with the collective numeral. In the EARLY ARCHAIC CHINESE texts in Dobson (1962) in which classifiers are optional, it may be noted that the classifiers for persons tends to be used with words for subordinates, slaves, and captives.

[21] CLASSICAL ARABIC substantives are conventionally cited in the numerative singular indefinite, usually ending in -un, I enclose this suffix in parentheses.

baqra or ra:s baqar and 'three cows' is either thala:th baqra:t or thala:thit rwa:s baqar. In OMANI then, the basic consideration is that a numeral cannot occur directly with a collective. Either a classifying noun, itself a non-collective, is interposed and is in direct syntactic construction with the numeral or the 'noun of unity' is used in its stead.

In CLASSICAL ARABIC as in OMANI there were collectives which did not have a noun of unity. Like other collectives they could not be governed directly by a number. The CLASSICAL grammarians prescribe that in such instances the preposition min 'from' must intervene. Thus with ʔibl a collective meaning 'camels' one had to say thala:thatu mina l ʔibli 'three of the camels'.[22] What has happened in OMANI and to a certain extent elsewhere is that the construction with min has been replaced by the use of a non-collective, we might say an individualized noun, as a classifier with the numeral while the collective follows as a kind of apposition.

The term "singulative" was first employed in CELTIC by Zeuss for the derivational formation which corresponds in these languages to the ARABIC "noun of unity". We may then talk of a "three term system" in such instances in which a collective which cannot be used with numerals is opposed to a singulative with its own singular and plural (or, in addition dual as in ARABIC). The plural of the singulative is thus distinct from the collective in such systems.[23]

In connection with the main thesis of the modelling of count nouns after mass nouns in quantitative constructions, it may be noted that there is an obvious analogy between mass nouns and collectives. In three term

---

[22] The example is from Gaudefroy-Demombynes (1952:372). He translates 'trois (individus) des chameaux'.

[23] It is of interest to note that such three-term systems also appear in NIGER-CONGO noun-class languages as in BANYUN and the MOMBAR dialect of SENUFO (Sauvageot, 1967) in which nouns may appear in three classes, a singular, a "limited" plural, and an "unlimited plural". With numerals only the former of these plurals may be used in BANYUN which is described by Sauvageot as "chiffrable" as distinguished from the unlimited plural which is "pas chiffrable". Sauvageot translates the unlimited plural of the word for a 'leaf' by a collective 'le feuillage'.

systems, the collective in addition to its central use in distinguishing genus from individual for organic species and human ethnic groupings, tends to be used also in some instances for nouns designating materials and even liquids in which case the singulative designates quasi-units in the sense described earlier. For instance, in CLASSICAL ARABIC there are examples such as khashab(un) 'wood' with khashabat(un) as its noun of unity meaning 'a piece of wood' and similar examples in modern dialects. In WELSH alongside dwfr 'water' one finds a singulative diferyn 'a drop of water'.

In CLASSICAL ARABIC there were two other systems besides that of the 'noun of unity' whose essential similarity with the noun of unity was noted by the grammarians. The noun of unity is only used with non-humans. For humans, in particular ethnic and occupational groups from a basic unmarked collective there was derived a singulative by suffixing -iyy(un) used in a manner exactly parallel to the noun of unity e.g. ru:m(un) 'Greeks (coll.)'; ru:miyy(un) 'a Greek', thala:thatu ru:miyyi:na 'three Greeks'. In addition the so-called nomina vicis (ʔismu'lmarrati) were derived from verbal nouns by the feminine singular suffix exactly as with the noun of unity to derive nouns designating individual acts. This also occurs in OMANI e.g. ḍʊḥk 'laughter', ḍʊḥka 'a laugh'.

A basically similar three term system is found in RUSSIAN in the period centering about the 16th century and has been described particularly by Unbegaun (1935). It developed on the basis of a Pan-Slavic collective formation in which yet another method of avoiding a direct construction between numerals and collectives had evolved, namely the use of a derivation of the numeral, the so-called collective numeral, governing the genitive of the collective.[24] In the RUSSIAN of the period under

[24] With this we may compare the CLASSICAL ARABIC construction with min 'from' cited above.

It is of interest to note that CLASSICAL ARABIC lexicons quote examples of raʔs 'head' and collectives without nouns of unity designating animals and root crops but with min intervening. Hence, this is a kind of transition between min + genitive prescribed by the Arab grammarians and ra:s + numerated noun of OMANI and other dialects. An example cited in Freytag (1930-5) is raʔsun mina'lkhəili lit. 'a head from the horses' in which khail(un) is a collective which also occurs in OMANI with ra:s (ra:s khe:l).

consideration there was once more a three-term system in which as an alternative to the more general SLAVIC use of the collective numeral, a singulative could occur with the ordinary (non-collective numerals). Most of the collectives were declined in the plural rather than the singular. Alongside these collectives there were singulatives either by derivation or sometimes with lexically distinct forms. As in MODERN RUSSIAN numbers larger than four governed the genitive plural in nominative constructions or nominative-accusative with inanimates, and in the other cases agreed with the plural in case. In these instances if the non-collective numerals were used, they required the plural of the singulative which in fact had no employment except in constructions with numbers. Where the collective was grammatically a singular as e.g. bratĭja 'brothers' (fem. sing.) the collective numeral could not be employed. As in ARABIC there was a special singulative suffix, in this case -in used for ethnic and occupational names.

For example, there was a collective with plural inflection krestĭjane 'peasants' (coll.) with a singulative krestĭjanin. For 'five peasants' in the nominative one could have either pjatero krestĭjan, the collective numeral with the genitive of the collective noun, or pjat' krestĭjaninov the ordinary (non-collective) numeral with the genitive plural of the singulative. However, as noted by Unbegaun, the first construction was uncommon. The most commonly employed was actually a third alternative for numerical constructions with personals and this corresponds to the use of an individualized classifying noun in OMANI. In this construction the noun čelověk 'person' occurred preceding the ordinary non-collective cardinal number, followed by the personal collective in the dependent genitive. Thus alongside, and in fact more often than the two alternative constructions given above, for expression 'five peasants' one could have pjat' čelověk krestĭjan. The word čelověk was itself a singulative corresponding to the collective ljudi 'people'. This relationship of course still

survives in contemporary RUSSIAN in that only čelověk may occur with numbers and its use in the plural is confined to numeral constructions.

To sum up, what is common to the RUSSIAN and ARABIC examples (and others e.g. CELTIC that might have been considered) is that, where there is a system of collectives, the direct construction of the numeral with a collective is avoided. Among the alternatives is the use of one or more non-collectives in construction with the numeral and more loosely joined syntactically to the collective which is in apposition or is a dependent (partitive) genitive.[25] In view of these and similar instances we may suggest as a hypothesis that when a language is an exception to the implicational universal that numeral classifiers imply the absence of compulsory plurals, what is involved is a subsystem of such a singular/plural language within which the basic opposition is collective/singulative rather than singular/plural.

A collective is somtimes defined as a noun which is grammatically singular but semantically plural. An example is Bielfeldt (1961: 296) in his grammar of OLD CHURCH SLAVIC who defines collective noun as follows. "Kollektivum – Subst., das in der grammatischen Form des Sg. eine Mehrheit von Gegenständen bezeichnet." However, the notion of collective in the ARABIC and RUSSIAN instances just considered and which is relevant to the present problem does not conform to this kind of formulation which seeks to define collective in terms of the categories of singular and plural, that is, as singular in form but plural in meaning.

Regarding the first of these criteria, singularity of form, we have seen that in sixteenth century RUSSIAN most of the collectives have plural inflection. It would seem in fact that the typical life history of the collective is that it starts out as a singular but with plural agreements or variation between singular and plural in more remote syntactic constructions

---

[25] A further method of individuation besides those mentioned in the text is examplified by IBIBIO (Kaufman, 1972) in which a phrasal compound of an individualizing noun plus the enumerated noun occurs, e.g. ákpɔ́ íffà kèèt 'stalk (of) firewood one'. Cf. ENGLISH "one rice grain"= 'one grain of rice'.

and tends to become a morphological plural in the course of time. This is an interesting topic which will not be pursued here.

In regard to meaning, if this is simply plural, then wherein does the distinction lie between the quantitative reference of the plural of the singulative and the collective? It would seem that the "true collective" is semantically neither singular nor plural. It is a transnumeral category which is neutral in respect to numbers as opposed to the singulative which involves countability or, as stated by Unbegaun (1935:262), it implies "l'opposition entre la collectivité et l'unité extraite de cette collectivité".

The generic noun (ʔismu ʾljansiyyu) of ARABIC, whose noun of unity is derived from it either by -at(un) as in tamr(un) 'dates (collective)', tamrat(un) 'a date' or by -iyy(un) as in ru:m 'Greeks (collective)', ru:miyy(un) 'a Greek', is excluded by the famous 13th century grammarian Ibn al-Ḥājib from his definition of plural since as he says "the expression is not constituted to express units but what contains the special quiddity (ma:hiyyat(un) lit. 'whatness') whether it be singular or plural". Concerning this passage, the commentator Raḍīuddīn says, "...to which we will add that the generic noun is applied to the few and the many.... So that if you eat a date, or two dates or deal with a Greek or two Greeks you may still say ʔakaltu ʾttamra ['I ate the dates (coll.)'] and ʕa:maltu ʾrru:ma ['I dealt with the Greeks (collect)'] whereas if they were plural this would not be allowable, as rija:l(un) [the ordinary broken plural of rajul(un) 'man'] is not applied to 'a man' or 'two men'."[26] The lack of relevance of specific number to collectives is also expressed by Maretić, a native speaker of SERBO-CROATIAN, a language with extensive and productive collective formations. In his grammar of SERBO-CROATIAN (1910:450) he says, "Therefore one cannot say, for example, deset kamenja ['ten stones (coll.)'] or petnaest perja ['fifteen feathers (coll.)'] etc., but instead deset kamena, petnaest pera [i.e. with the gen. plural singulative], because when someone mentions a definite number, he then thinks of individual things; but for

[26] These passages are quoted from Howell (1880-1900: 1054-5).

that which forms a collectivity [što je u hrpi, lit. 'is in a heap'] the number is not known."

It should be noted that in ENGLISH the most commonly cited example of a numeral classifier construction 'head of cattle' involves a collective. There exists in ENGLISH what might be called a miniature system of collective/singulative e.g. Irish//Irishman/Irishmen, police//policeman/policemen. I have tried the following sentence on a number of native speakers of ENGLISH. "Last night I was picked up by the police." They all denied that they would not be surprised to learn to that only one policeman was involved. My own reaction to constructions of numbers with these collectives is that small numbers seem definitely ungrammatical but fairly large numbers seem fairly natural. A meeting of twenty police or one hundred faculty seems acceptable but the phrase ' a meeting of three police' is definitely strange.[27]

The development of the construction čelověk in RUSSIAN with collectives and the fairly extensive system found in OMANI and elsewhere in ARABIC suggests that classifiers in the large majority of classifier languages without plural inflections are performing the same individualizing function as both classifiers and singulative affixes in languages with collectives. We should expect then that in the typical classifier language, the classifiable noun when not accompanied by a classifier should show the same lack of numerical determination that we have found with collectives in languages like ARABIC.

Emeneau (1951: 85) describes the VIETNAMESE noun when unaccompanied by a classifier in terms quite remeniscent of Raḍīuddīn in regard to the generic noun of ARABIC:

---

[27] Actually, there exists a variety of uses of collectives which should be distinguished. These are heterogeneous and all that they have in common is that they do not involve the results of counting. These uses include true generic uses in generalizations, "hypothetical" uses in such sentences as 'He went hunting for deer.' and references to actual collections which are either so large as not to be practically countable or potentially countable but not actually counted, or counted where the numerical result is irrelevant. This topic is not pursued in this paper.

> "A non-numerated substantive phrase ... lacks any indication of number or individuation; that is when there is no explicit indication of number, a number is entirely free of reference to the number category. For example, tôi muốn mua sách 'I want to buy book(s)'. There is absolutely no indication how many books are intended."

A considerable number of classifier languages (e.g. many IRANIAN and TURKIC languages, KOREAN) have what are generally described as plural affixes. However, closer examination seems to show that in almost every instance the "unmarked" singular is in fact a form which, like the collective in languages with a compulsory plural, is non-committal in regard to number.[28] For an explicit statement to this effect, parallel to those in regard to the other linguistic types we have been considering, reference may be made to Kononov (1960:75) who states concerning UZBEK that words like 'girl' and 'bird' without any grammatical indication do not contain any indication of number. They represent an undivided (nečlenimoje) totality. When the suffix of plurality -lar is added they became a totality consisting of individual members (členimoje). What is hypothesized then, is that in the usual classifier language (i.e. without inflection for number), classifiable nouns in their isolated form, that is when not accompanied by a classifier or a plural marker, are like collectives in their semantic non-specification of number and in their avoidance of a direct number construction. The classifier is an individualizer which performs the same function as a singulative derivational affix in languages with the collective/singulative opposition.

In two grammatical descriptions of classifier languages I have found a point of view similar to the one expressed here. One is Dobson's work on EARLY ARCHAIC CHINESE in which he states (1962:28):

> "it is not a feature of 'substantival quality' that it distinguishes class and member, between the genus itself and 'an instance of' or instances of .... In EAC certain of the distinctions are made when a noun occurs in a syntagma in which the elements are distributed as enumerated noun/number/quantification...."

By quantification is here indicated what is usually called a classifier.

---

[28] Possible exceptions include OSSETE, PASHTO and TLINGIT.

Another statement is that of Grjunberg (1963:46) in his grammar of NORTH AZERBAIJANIAN TAT, an IRANIAN language:

> "As has been already indicated, what is formally in Tat the singular is the expression of an undivided multitude (množestvo) of objects and almost always has a particular kind of collective meaning .... In order to supply such a substantive expression with a quantitative meaning by means of a number, it is first necessary to select a unit for counting. For this reason the numeral does not usually stand immediately before the substantive. Between them one places a word indicating such a unit of calculation."

If the general point of view expressed here is taken as at least a working hypothesis, one of the further problems to be considered is the following. It was seen that in ARABIC and RUSSIAN, the use of a noun as a classifier was but one of a series of functional alternatives such as a derived singulative of the enumerated noun or a special derived form of the numeral. One might conjecture that the choice of a noun itself as an individuator rests on the fact that, as a general rule, these languages have a very weak or even non-existent development of derivation. In general they use syntagmatic structures consisting of full words. This, it would seem, is what is meant by the traditional notion of isolating languages as a type.

We have seen what might be called, anthropomorphically, the aversion of collectives to direct construction with a numeral and the intervention of an individuated noun, the classifier, as one of the devices to avoid this direct confrontation. This aversion has, therefore, as its natural counterpart, the corresponding attraction to the classifier and an immediate constituent structure in which the numeral goes directly with the classifier while the numeral+classifier combination as a whole enters into a more remote construction with the enumerated noun. In languages with substantival inflection for number and a singulative/collective opposition it was noted that the numeral governs the classifier in respect to such categories as number and case, while the enumerated noun is in apposition to or stands in an adnominal construction (essentially partitive) to this combination.

This arrangement underlies a number of synchronic generalizations that may be made regarding classifier languages proper. We may summarize these as follows.

1) Of the six possible word orders among the three elements Q (quantifier), Cl (classifier), N (enumerated noun), only four occur -- 1. Q-Cl-N; 2. N-Q-Cl; 3. Cl-Q-N; 4. N-Cl-Q.[29] The two non-occuring orders Cl-N-Q and Q-N-Cl have the property that the quantifier and the classifier are separated by the head noun.

2) There is frequent variation within languages between orders 1 and 2 or between 3 and 4. In other words the relative order of quantifiers and classifier remains unchanged but the combination of the two may vary between placement before or after the head noun. The rare variation between Q-Cl and Cl-Q is of three kinds. In BODO, a SINO-TIBETAN language there are two distinct subsystems, the indigenous (Cl-Q) and that borrowed from ASSAMESE (Q-Cl). In BENGALI, according to Chatterji (1926:777), the usual order Q-Cl may be reversed to express numerical approximation. In most THAI languages the Q-Cl order generally holds but the order with the number 'one' is Cl-Q.

3) The connection between the numeral and the classifier is so close prosodically that they may have one accent, in which case it is on the numeral and there may be fused forms such that analysis becomes difficult. In this case, e.g. the MICRONESIAN languages, the numerals are generally said to form a number of series. In many languages, analysts consider the numeral + classifier construction to be a single word.

4) The Q-Cl combination may often be separated in certain constructions from the enumerated noun.

5) The anaphoric construction of Q-Cl without overt expression of the noun occurs in all of these languages.

---

[29] Q for 'quantifier' is used here because not only numerals but also the numerical interrogative 'how many?' and less frequently indefinite quantifiers such as 'few', 'many' occur in the same position as the numeral in classifier constructions. An apparent exception is IBIBIO with Cl-Q-N (see footnote 25 above). Note, however, that the numeral here is really in construction not with the enumerated noun as such but with the phrasal compound.

These facts and the general structuring which they exhibit have a bearing on the question of the interpretation of those instances cited in an earlier section of this paper of nouns which commonly appear without classifiers. These include words like 'day', 'month', 'time' (in phrases corresponding to ENGLISH 'three times'), 'foot', 'mile', and currency expressions. As can be seen, these can all be interpreted as measures. In addition the words for 'person' and 'thing' may sometimes occur both with other items as classifiers of very general scope but also in their lexical meaning without previous mention in the context (e.g. for 'person' in many THAI languages). In these instances then, where one might have expected 'person' to act as its own classifier we merely find the translational equivalent "three person" rather than repetition e.g. "person three-person".

Two interpretations of these phenomena have been offered by writers of grammars sometimes regarding the same language. For example Burling (1961: 266) in his grammar of GARO analyzes these words as classifiers without head nouns while noting that in previous grammars of the language they were analyzed as head nouns without classifiers. Burling's analysis would seem to be the more justified. In such instances the same close syntactic construction between numeral and "unclassified noun" is formed as between the numeral and the numeral classifier in the tripartite construction. CUNA (Scherzer, forthcoming) provides here a particularly striking confirmation. As often elsewhere the classifier forms a single word with the numeral while the head noun is separate. This is shown here by the stress pattern e.g. óme wár-po 'woman classifier-two'. In expression such as 'one day', 'day' is seen to be a classifier from the stress pattern ípá-k^{w}en 'day one' rather than *ípa k^{w}én.

Moreover, where the word order is noun + numeral classifier, these phrases invariably have numeral + "unclassified noun" rather than the opposite order and similarly where the tripartite order is classifier + numeral + noun the order of these phrases is "unclassified" noun + numeral. It could be maintained that in measure phrases, the place of the head

noun is essentially taken by a verb; it is the verbal action that is being quantified. For example, "he traveled two days" is equivalent to "he performed two days of travel".[30] There are both measure and count verbal constructions. For most languages a single lexical item, the equivalent of the ENGLISH word 'time', is used with numerals or a special set of numeral is used (e.g. LATIN), but another alternative which is found, for example, in ARABIC, is the "cognate" verbal noun meaning a single instance of an act which may then cooccur with numerals. It was noted earlier in passing that in ARABIC the so-called maṣdar or "infinitive" is a verbal noun which may then take the same feminine suffix of the noun of unity when subject to count construction and that the ARABIC grammarians noted the essential parallelism, collective noun : noun of unity = verbal noun (maṣdar) : nomen vicis. Similar phenomena occur in languages without the collective/singulative contrast. In BODO there are examples such as pay-tam pay "comings-three came". There is also the use of nouns which are neither general for all verbs like 'time' in ENGLISH nor cognate verbal nouns e.g. MANDARIN kànle liáng-ǰyen "locked two eyes" 'looked twice'.

The logical possibility exists, then, that a language might have a system of verbal classifiers each of which would be used with a particular class of verbs and an accompanying numeral. However, this possibility never seems to be realized in the systematic way in which it so often is for nouns. The distinction between mass and count then applies to verbal action and is related to aspect. Durative : punctual = measure : count, 'He has been laughing for two minutes.' versus 'He laughed twice.'[31] There is the widespread phenomenon of "plural verbs" marking plural

---

[30] A few linguists have used such terms as verbal measures, e.g. Smalley (1961).

[31] For a discussion of the mass/count distinction in relation to verbs cf. Leech (1969:134-5). He notes that "...not only noun meanings but verb meanings can include the factor 'countability' ".

action as against a single act. Once again the abstract possibility of incorporating count distinction beyond the singular/plural dichotomy in the verb exists, e.g. a verbal form meaning 'to perform X three times' which does not seem to occur anywhere.

The "attraction" of the individuated noun to numerical expression as contrasted with the "aversion" of the collective noun, has already been discussed. The relation of measures to numerical expressions which we have hypothesized as a model for count construction is in a sense even closer than that of the individuated noun and is somewhat different in nature. They are syncategorematic with quantifiers in that they have no reality without them. A word like 'ounce' when used,not merely mentioned,has its raison d'être in being accompanied by a quantitative expression. Only perhaps in metalinguistic discourse e.g. ounce is a measure noun or the Imperial gallon is larger than the usual American gallon can it be abstracted from quantity. Ounces are not counted like apples. If I have a set of six apples, I can ask about physical characteristics of the apples in abstraction from their number, e.g. their color, but not so with six ounces. I can imagine a large finite number of all past, present and future apples but I cannot number ounces in the same way. Similarly, when a physical object like a cup is being used as a measure, three cups full of tea is different from three cups of tea . I might indeed use the same cup three times. Not being physical objects they are not susceptible of the distinction between collective and individual. A further example of the contrast between abstract measures and concrete objects is the difference between monetary value and actual coins or bills. Twenth-five cents and a quarter as a coin are not the same thing.

This is perhaps why measures in many languages with inflected singular and plural tend to use measures in the unmarked singular. Note the distinction in German between zehn Pfennig as an amount and zehn Pfennige with the plural of the noun as ten coins each worth one Pfennig.

This brings us to a final consideration regarding substantival number in the numeral classifier constructions. As one might expect in languages

of this kind with singular and plural the number frequently has the same construction with classifiers as it does with other nouns, e.g. plural with 3-10 in OMANI ARABIC. However, in classifier languages without inflectional plurals neither counters nor measures ever take plural markers and, unlike typical noun class systems, the classifiers themselves practically never vary lexically for number. The only exception I have encountered is in GUAYMI, a CHIBCHAN language (Alphonse 1956:13), in which i 'person' is used as a classifier with 'one' and ni 'people' with numbers greater than one.

In languages in which the demonstrative occurs with classifiers there seems invariably to be a single 'plural' classifier replacing the ordinary classifier but only in the demonstrative construction, not in the numerical construction. For example in MANDARIN the classifier bĕn required with shu 'book' with any number, e.g. i bĕn shu 'one book', san bĕn shu 'three books' occurs with the demonstrative also, chè bĕn shu 'this book' but only in the singular. For all nouns, the ordinary classifier is replaced by hsie 'some' to form the plural with demonstratives, chè hsie shu 'these books'. Basically similar are the BENGALI, ASSAMESE, etc. 'definitives' which are suffixed to the noun to make them definite as well as occurring with the numerals as classifiers, e.g. BENGALI pāṅc-khana boi "five flat-object book", boi-khana 'the book'. In the plural definite all classifiers are replaced by the plural gulo, e.g. boi-gulo 'the books'. This plural marker cannot occur with numbers. Unit counters, then, behave very much like the measures which have been hypothesized here as their models. The notion of 'unit' seems to take on this same abstractness which characterizes measures and tends to make them take an unmarked invariable form with numbers. For the counters, whose lexical source is generally transparent, are like the singulative in containing two semantic moments, the concrete lexical meaning 'head', 'piece', 'grain', or whatever it may be, and the notion of 'unit of counting' as such. It is evidently the latter that tends to assume the same abstractness of meaning that is inherent in measures.

Finally, in order to place the present paper within the framework of the more general study of which it is a part, four other basic topics will be briefly considered and in some instances hypotheses will be outlined.

1) From the fact that certain languages have developed the numeral classifier system, it by no means follows that it must have appeared in a single step in all numerical constructions and compulsorily. There is some evidence that it tends to appear first as focus particularly in answers to quantitative WH-questions and later spreads to other constructions.

There is indeed a general problem here in relation to the main thesis of this paper. If some method of individualization with specific quantity is required where the noun has a general unmarked form, how is it that languages may have certain syntactic constructions requiring this and others not, and how can the construction be optional in other languages? This is part of the broader problem as to why the implicational relation between classifiers and lack of compulsory number in the substantive is just that and not a mutual implication, that is, a logical equivalence. There seem to be languages without compulsory number inflection, which likewise have measure constructions and yet do not have classifiers. In other words, as explained in the initial section of this paper, we have at best necessary but not sufficient conditions.

2) Another basic problem relates to the lexical sources of classifiers and their semantic relation to the head nouns. There seem here to be three main types: A) superordinate terms such as 'person' as a classifier for humans and 'tree' for individual 'species'; B) items in one-to-one relation to the objects being counted, among the most common of these are 'head' for animates and 'trunk' or 'stalk' for trees; C) words which themselves designate arbitrary or insignificant units like 'piece', 'grain', etc. It was seen earlier that these exist quite generally in languages which have measures and having somewhat equivocal status, they are capable of spreading semantically to structured items. For non-structured units these terms often relate to the verbal action required to produce them, analogous to ENGLISH slice in slices of bread.

3) This last example brings up a further major area of investigation, the semantic changes of the classifiers in terms of changes in the nouns that they classify. Here a thoroughgoing comparison with semantic changes of class indicators in other types of numeral classification systems is of value. These processes are in many respects similar to those involved in ordinary lexical semantic change. However, they are, so to speak, more unrestrained in their capacity of generalization because in the vast majority of instances they are semantically redundant. The rôle of shape in classification has been singled out for particular attention by some analysts. It is indeed a recurrent phenomenon that we find classifiers which cooccur with groups of nouns which have as their common semantic feature one of the following: a) long narrow object (one-dimensional), often subdivided into cylindrical and non-cylindrical; b) flat object (two dimensional); c) round object (three-dimensional). This latter tends to include large bulky objects of whatever shape. These classifiers apply primarily to inanimates but they sometimes include various categories of animates. For example 'snakes' or larger quadrupeds are often classified as long, narrow objects.

Insofar as classification is applied basically to countable, concrete objects it is not difficult to see that semantic criteria of shape provide the broadest possibilities for generalization as being that which otherwise heterogeneous physical objects have in common. In many instances the same lexical item used as a classifier has in diverse languages become one of the basic shape classifiers, notably 'stalk' or 'trunk', an item in one-to-one relation with plants and trees, for long narrow objects and 'grain', a "quasi-unit classifier" for round objects.

The frequent occurrence of what is sometimes called the general classifier is to be interpreted in dynamic terms as the ultimate result of semantic generalization of one of the widespread classifiers, generally one of the shape classifiers and most typically the round object classifier to the point at which it not only itself cooccurs with a very large and heterogeneous group of nouns, but may be used as an alternative to almost any

other classifier. There is evidence in some instances regarding the diachronic expansion of these classifiers. Often the spread of such classifiers is confined to inanimates. For example, in regard to VIETNAMESE, Thompson (1965:196) notes that "In modern VIETNAMESE the general classifier cái is coming to be used more and more at the expense of other specific classifiers, especially with nouns denoting inanimate objects which in traditional usage go with one of the rarer classifiers ...." In other instances, e.g. MANDARIN gè, the general classifier is used also with persons.

There is an enlightening parallel here with the process of consolidation and simplification found in noun-class languages. In NIGER-CONGO languages there is a tendency for one of the non-personal classes to become the "general" class paralleling in its semantics that of the general classifier in heterogeneity of meaning, statistical frequency and tendency to be used in place of other non-personal classes.

A similar phenomenon is found in what might be called possessive classificational systems. In many OCEANIC and AMERIND languages the very common contrast between intimate and non-intimate possession has been elaborated through the split of the latter into classes based on the use of various classifying nouns which takes the possessive affixes in place of the noun designating the possessed item which is then placed in apposition. For example, in MATACO, a language of the CHACO, 'my dog' is, literally "my-animal dog" and 'my house' is "my-property house". Such systems also tend to develop a "general" class. For example, in SONSOROL, a language of Micronesia which like many OCEANIC languages has simultaneously possessive and numeral classifier systems which are independent of each other, one of the noun bases of the possessive system ja- is described as signifying "general possession, not covered by any other class" (Capell 1948:13).

4) One of the lines of development of such systems is by syntactic spread to other constructions than the numeral classifier construction

proper. The synchronic universal seems to hold that whenever a numeral classifier construction is also used in non-quantifier constructions, the construction with demonstratives is one of these, often the only one. The use of numeral classifiers with demonstratives occurs in a number of geographically separate areas and some of these instances at least must be historically independent, e.g. THAI, VIETNAMESE, MODERN CHINESE, BENGALI, NAURU (Micronesia) and KIRIWINA (Trobriand Islands east of New Guinea and geographically distant from Nauru). Demonstratives would seem to have, like numbers, a special relation to individuated non-collective expressions but the details of this process remain to be investigated.

Throughout this paper I have emphasized the tentativeness of the conclusions advanced and that it is to be viewed more in the light of a progress report than a definitive statement. Its value, it may be hoped, is to show that the method of dynamic typological comparison can help in investigating significant problems which have, on the whole, not been discussed very much in recent linguistics, and can also, by the consideration of empirical linguistic data from a great variety of languages at least, open the possibility of rational solutions to such traditional problems as the origin of gender and noun classificational systems in general.

BIBLIOGRAPHY

Alphonse, E.S. 1956. Guaymí grammar and dictionary. Washington.

Bataillon, M. 1932. Langue d'uvea. Paris.

Bielfeldt, H.H. 1961. Altslawische Grammatik. Halle.

Brockelmann, C. 1908-13. Grundriss der vergleichenden Grammatik der semitischen Sprachen. Berlin.

Burling, R. 1961. A Garo grammar. Poona.

_____. 1965. "How to choose a Burmese numeral classifier", in Context and meaning in cultural anthropology (Festschrift Hallowell), ed. by M. Spiro, 243-264. New York.

Cappell, A. 1948. Grammar and vocabulary of the Sonsoral-Tobi language. Washington, D.C.

Chao, Y.R. 1968. A grammar of spoken Chinese. Berkeley and Los Angeles.

Chatterji, S.K. 1926. The origin and development of the Bengali language. Calcutta.

Dobson, W.A.C.H. 1962. Early Archaic Chinese. Toronto.

Emeneau, Murray B. 1951. Studies in Vietnamese (Annamese) grammar. Berkeley and Los Angeles.

_____. 1956. India as a linguistic area. Language 32, 3-16.

Ferguson, C.F. 1964. "The basic grammatical categories of Bengali", in Proceedings of the 9th International Congress of Linguists, Cambridge, ed. by Horace Lunt, 881-890. The Hague.

Freytag, G.V. 1830-5. Lexicon Arabico-Latinum. Berlin.

Gaudefroy-Demombynes, M. 1952. Grammaire de l'arabe classique. (3rd rev. ed.) Paris.

Goldschmidt, W. 1951. Ethics and the structure of society: an ethnological contribution to the sociology of knowledge. American Anthropologist 53, 506-524.

Greenberg, Joseph H. 1966. Synchronic and diachronic universals in phonology. Language 42, 508-517.

_____. 1969. Some methods of dynamic comparison in linguistics, in Substance and structure of Language, ed. by J. Puhvel, 147-203. Berkeley and Los Angeles.

_____. 1970(a). Some generalizations concerning glottalic consonants, especially implosives. International Journal of American Linguistics 36, 123-45.

_____. 1970(b). The role of typology in the development of a scientific linguistics, in Theoretical problems of typology and Northern Eurasian languages, ed. by L. Dezsö and P. Hajdú, 11-24. Budapest.

Grjunberg, A.L. 1963. Jazyk severoazerbajdzhanskix tatov. Leningrad.

Howell, M.S. 1880-1900. A grammar of the Classical Arabic language. Allahabad.

Jacob, J.M. 1965. Notes on the numerals and numerical coefficients in Old, Middle and Modern Khmer. Lingua 15, 143-162.

Kaufman, E.M. 1972. Ibibio dictionary. Office of Education. Washington, D.C.

Kononov, A.N. 1960. Grammatika sovremennogo uzbekskogo literaturnogo jazyka. Moscow-Leningrad.

Leech, G.N. 1969. Towards a semantic description of English. Bloomington and London.

Maretić, T. 1910. Grammatika i stilistika hrvatskoga ili srpskoga književnoga jezika. Zagreb.

Mitchell, T.F. 1956. An introduction to Egyptian Colloquial Arabic. Oxford.

Noss, R.B. 1964. Thai reference grammar. Washington.

Reichardt, C. 1894. Ein arabischer dialekt gesprochen in Oman und Zanzibar. Berlin.

Sanches, M. 1971. Numeral classifiers and plural marking: an implicational universal. (Unpublished manuscript).

Sauvageot, S. 1967. Note sur la classification nominale en bainouk, in La classification nominale dan les langues Nègre- Africaines, 225-236. Paris.

Scherzer, J. (forthcoming). Cuna numeral classifiers. To appear in A.A. Hill Festschrift.

Sen-Gupta, S. 1970. Genesis of the numeral classifier and its function as a morphological category. Proceedings of the 10th International Congress of Linguists, Vol. 3, 675-82. Bucarest.

Smalley, W.A. 1961. Outline of Khmu structure. New Haven.

Thompson, L.C. 1965. A Vietnamese grammar. Seattle.

Unbegaun, B. 1935. La langue russe au XVIe siècle (1500-1550) I, la flexion des noms. Paris.

Vinogradov, V.V. 1934. Očerki po istorii russkogo literaturnogo jazyka XVII-XIX vv. Moscow.

Whorf, B.L. 1941. The relation of habitual thought and behavior to language, in Language, culture and personality, essays in Memory of Edward Sapir, ed. by L. Spier et al., 75-93. Menasha.

Windstedt, R. 1945. Colloquial Malay. Revised edition. London.

_____. 1957. An unabridged Malay-English dictionary. 2nd edition. Singapore.